The statistical method used in the prospective studies have been reviewed by Brownlee (Am Stat Assoc J 60: 722-738, 1965). This article is so important that it is reproduced in its entirety.

A REVIEW OF "SMOKING AND HEALTH"*†

K. A. Brownlee
University of Chicago

Although there were earlier papers in the medical literature, e.g. Müller [1], the possibility of an association between cigarette smoking and lung cancer first became common knowledge among statisticians generally with Doll and Hill's [2] retrospective paper published in 1952. A retrospective study is one in which a group of, e.g., cancer patients is classified by their smoking habits, and a second group, the controls, known to not have cancer, but sampled from the same population as the cancer patients, is also classified by their smoking habits. The resulting 2×2 table, in which in principle the two sample sizes are fixed, can be tested for independence. The evidence for the association became more convincing with the publication in 1954 of two reports on prospective studies, one in England by Doll and Hill [3] and one in the United States by Hammond and Horn [4]. In a prospective study, a population is sampled and then each individual is classified by, e.g., his smoking habits and by his experience with respect to cancer after an interval of time sufficient to allow a reasonable number of cases of the disease to develop.

I have observed amongst some statisticians a wistfulness that statistics has not so far played a larger part in science generally. Since this association between smoking and lung cancer, interpreted by many as one of causation, is of prime importance, one might have expected it to be greeted with enthusiasm and loud admiration. On the contrary, the comments and reactions of the statistical profession have been very restrained. By and large, in fact, the silence has been deafening.

In 1955, Neyman [5] pointed out a possible fallacy in retrospective studies, but carefully asserted in a footnote that he had no direct information on the association between smoking and cancer. Incidentally, this footnote contains the curious remark:

"A referee warns me that in spite of the fictitiousness of the figures in Table 1 and in spite of the emphasis on the methodological character of my remarks, the 'tobacco people' may pick up the argument and use it for publicity purposes."

Also in 1955, Berkson [6] and Mainland and Herrera [7] pointed out certain types of bias which could affect prospective studies.

In 1955, Cutler [8] reviewed the literature. His overall conclusion at that time was "There is disagreement whether the evidence at hand warrants a conclusion that smoking and lung cancer are causally related." He also remarked "If lung cancer should consistently occur more frequently in smokers than in non-smokers, in various subgroups of the populations studied, the cas

* An invited review article on "Smoking and Health," Report of the Advisory Committee to the Surgeon General of the Public Health Service, U. S. Department of Health, Education, and Welfare. Public Health Service Publication No. 1103. Superintendent of Documents, Government Printing Office, Washington, D. C. 20402. xvii, 387 pp. $1.25. Paper.

† I wish to thank J. Berkson, L. A. Goodman, G. W. Haggstrom, and W. H. Kruskal for comments they made on an earlier draft of this review. I also wish to thank several referees for their suggestions.

for a cause-effect relationship will be greatly strengthened. To refute effectively the hypothesis that smoking is a cause of lung cancer would then require a reasonable explanation, other than causation, for the consistently observed association between smoking and lung cancer."

In 1957, Sir Ronald Fisher [9] offered such an explanation: "that cigarette smoking and lung cancer, though not mutually causative, are both influenced by a common cause, in this case the individual genotype."

Berkson in a series of papers, e.g. [10] and [11], expressed doubt regarding the demonstration of causation. He emphasized that a higher death rate was shown for almost all diseases, and that if this is accepted as reflecting the excess deaths due to smoking then about 40 per cent of all deaths among cigarette smokers must be attributed to their smoking.

In 1959 Cornfield et al [12] concluded that "the consistency of all the epidemiologic and experimental evidence also supports the conclusion of a causal relationship (of lung cancer) with cigarette smoking, while there are serious inconsistencies in reconciling the evidence with other hypotheses which have been advanced."

Of the comments of statisticians, the above seem to be among the more important.

In 1962 the Surgeon General of the Public Health Service of the U. S. Department of Health, Education, and Welfare moved to appoint a committee to "assess available knowledge in this area [smoking vs. health] and make appropriate recommendations" (page 7). It is stated (page 8) that the function of this committee was to make "an objective assessment of the nature and magnitude of the health hazard" and "this committee would produce and submit to the Surgeon General a technical report containing evaluations and conclusions." However, "Recommendations for action were not to be a part" of the committee's responsibility.

The formation of the committee was a formidable operation. The Surgeon General met on July 24, 1962, with representatives of the American Cancer Society, the American College of Chest Physicians, the American Heart Association, the American Medical Association, The Tobacco Institute, Inc., the Food and Drug Administration, the National Tuberculosis Association, the Federal Trade Commission, and the President's Office of Science and Technology. It is noteworthy that no statistical society was represented. Also the list of organisations appears to be heavily weighted towards government agencies and organisations large, general, and active in public relations, and to have low representation of societies with specifically scientific outlooks. The participants of the July 24 meeting compiled a list of 150 scientists and physicians. This list was then screened by these participants, each organisation represented having a power of veto. From the intersection of all the sets of non-vetoes (which might, but apparently did not, have turned out to be the empty set, though we are not told its size) ten were selected for the Committee. Of the ten, only one is listed as a statistician. Since a substantial fraction of the relevant issues are statistical, one might question whether a one in ten representation for statistics was sufficient. However, one further member of the Committee is listed as an epidemiologist.

The committee met for nine sessions of from two to four days between November 1962 and December 1963, and in addition there were "uncounted" meetings of subcommittees. The committee had the assistance of a professional staff of eight, with a secretarial and technical staff of fifteen. The committee lists 189 individuals or organisations who made "contributions." The literature surveyed was enormous. A basic bibliography of more than 6000 articles from 1200 journals up to 1959 was supplemented by an additional 1100 titles provided by the National Library of Medicine. The Report lists in the references to its various chapters about 900 references. Some of the individual studies considered were themselves enormous. For Hammond's latest report [13], 68,116 volunteers enrolled 1,078,894 men and women.

The Report represents a tremendous sifting, by a large group, of an enormous literature into 387 pages, and this review will confine itself to a brief summary of its contents and an assessment of its major conclusions, with emphasis on statistical and methodological aspects.

The front matter contains no trace of a statement of date of publication. The very last page (page 387) does have as a footnote the notation "U. S. Government Printing Office: 1964 0-714-422" in which I presume the "1964" is the year of publication.

Chapter 1 reviews the evolution of interest in the possible deleterious effects of smoking and the procedures for the establishment of the Committee. Chapter 2 describes the operation of the Committee through its subcommittees, consultants, etc. Chapter 3, "Criteria for Judgment," discusses "the Epidemiologic Method." It remarks (pages 20–21) "Statistical methods cannot establish proof of a causal relationship in an association. The causal significance of an association is a matter of judgment . . . To judge or evaluate the causal significance of the association between the attribute or agent and the disease, or effect upon health, a number of criteria must be utilized, no one of which is an all-sufficient basis for judgment. These criteria include:

(a) The consistency of the association
(b) The strength of the association
(c) The specificity of the association
(d) The temporal relationship of the association
(e) The coherence of the association"

The Report does not define these terms at this point, though they are discussed in pages 182–189 in connection with lung cancer. The Report does not state whether these are jointly sufficient. Nor does this discussion of "the epidemiologic method" give any mention of the factor of self-selection, as discussed by Yerushalmy [14], Yerushalmy and Palmer [15]. Later in the Report (pages 180–181), there is a mention of selection bias, both by the operator of the survey and by the individual whose cooperation is being sought. The subsequent discussion (page 181, last paragraph) is concerned with the results of selection of the type caused by exclusion, or decreased probability of inclusion of persons sick or about to be sick: the discussion does not bear on the well-known fact that volunteers (or equivalently, cooperative persons) may, and often do, differ in their disease experience.

Chapter 3 also has a section on "Causality," (pages 20–21), the essential part of which appears to be "It is recognized that often the coexistence of several factors is required for the occurrence of a disease, and that one of the factors may play a determinant role, i.e. without it the other factors . . . are impotent . . .

The word *cause* is the one in general usage . . . and is capable of conveying the notion of a significant, effectual, relationship between an agent and an associated disorder or disease in the host."

Presumably in the lung cancer situation *cause* must be defined probabilistically. A very over-simplified sketch might be as follows. Suppose that each individual and his circumstances and environment can be classified dichotomously on each of n classifications, so that he is either an E_1 or an E_1^c, an E_2 or an E_2^c, etc., where E_1^c is the complement of E_1, etc. Suppose now that

$$\Pr\{E_1 | E_2 E_j \cdots E_k\} \neq \Pr\{E_1 | E_2^c E_j \cdots E_k\},$$

where $E_j \cdots E_k$ is some specified subset of conditions $E_3, \cdots, E_n$. Then, for that set of conditions $E_j \cdots E_k$, the probability of E_1 given E_2 is different from the probability of E_1 given E_2^c. Since the presence or absence of E_2 affects the probability of E_1, it seems reasonable to say that E_2 is a *cause* of E_1. The practical significance of E_2 as a cause of E_1 will depend on two conditions:

(a) If the left hand side of the above inequality, though not exactly equal to the right hand side, is numerically quite close, then for that set of $E_j \cdots E_k$ the causative association may be of small practical importance.

(b) If the inequality holds only for, say, one particular subset $E_j, \cdots, E_k$, and for all other subsets equality holds, and if the subset $E_j, \cdots, E_k$ occurs in the population with low probability, then $\Pr\{E_1 | E_2\}$, while not strictly equal to $\Pr\{E_1 | E_2^c\}$, will be numerically close to it, and then E_2 as a cause of E_1 may be of small practical importance. These considerations are related to the Committee's responsibility for assessment of the *magnitude* of the health hazard (page 8). Further complexities arise when we distinguish between cases in which one of the required secondary conditions $E_j, \cdots, E_k$ is, on the one hand, presumably controllable by the individual, e.g. the eating of parsnips, or uncontrollable, e.g. the presence of some genetic property. In the latter case, it further makes a difference whether the genetic property is identifiable or non-identifiable: for example, it could be brown eyes which is the significant subsidiary condition E_j, and we could tell everybody with not-brown eyes it was safe for *them* to smoke.

Chapter 4 is a set of Summaries and Conclusions. Chapter 5 gives consumption of tobacco products, including chewing tobacco and snuff, for the United States for selected years from 1900 to 1962. Chapter 6 is on "Chemical and Physical Characteristics of Tobacco and Tobacco Smoke." Chapter 7, "Pharmacology and Toxicology of Nicotine," concludes that nicotine is unlikely to be important.

In Chapter 8 we come to a discussion of the relationship between smoking and mortality. Subsequent chapters deal with specific groups of diseases: 9, Cancer; 10, Non-Neoplastic Respiratory Diseases; 11, Cardiovascular Diseases; 12, Other Conditions. Chapter 13, "Characterization of the Tobacco

Habit," concludes that smoking is a *habituation* rather than an *addiction*, "since once established there is little tendency to increase the dose," etc. (page 354). Chapter 14, "Psycho-Social Aspects of Smoking," states that "The overwhelming evidence points to the conclusion that smoking is to a large extent psychologically and socially determined" (page 377). The "psychologically" part of this conclusion is apparently based on Eysenck [16] who found that cigarette smoking was positively associated with extroversion, and that both non-smokers and pipe smokers (the study was made in England, so presumably there were too few cigar smokers to consider) were less extroverted.

The evidence for the "socially" part of the conclusion seems in its summarized form to be not very quantitative. The Report forms the conclusion that "white collar professional, managerial and technical occupations contain fewer smokers than craftsmen, salespersons, and laborers" (page 363), but does not give any figures. However, "As to separate class-linked variables, income does not seem to be related in a consistent manner to prevalence of smoking." The Report finds "The relationship between smoking and education is unclear," and "The proportion of smokers is roughly the same among whites and non-whites." Men smoke more than women, though the discrepancy has been decreasing in recent years. Also, "Smoking (of any kind) is most prevalent among the divorced and widowed and least among those who have never been married, except that among persons over 45, never married are as likely to be smokers as the married" (page 364). Some religions affect smoking, and the rural farm population smoke less than the rural non-farm population. There is little difference between the latter and the urban population. Few figures are quoted in this discussion, so one cannot assess how important these various factors are.

Chapter 15 deals with the morphological constitution of smokers. What seems the most useful information is from Seltzer [17], who in 1942 obtained anthropometric measurements on 922 Harvard undergraduates and in 1950 ascertained their smoking habits (with an 81 per cent response). He found that cigarette only smokers were statistically larger in a number of anthropometric indices than non-smokers, averaging 4.37 pounds heavier, pipe only smokers larger still, and cigar smokers the largest. The Report concludes, however, "The available evidence suggests the existence of some morphological differences between smokers and non-smokers, but is too meagre to permit a conclusion" (page 387). The Report does not give reasons for declining to accept Seltzer's results.

We turn now to a detailed consideration of Chapters 8 and 9.

Chapter 8 reviews the seven major prospective studies and concludes that the overall death rate for cigarette smokers is about 70 per cent higher than for non-smokers. For cigar and pipe smokers the effect was minor. The mortality ratio, i.e. the ratio of the death rate for cigarette smokers to the death rate for non-smokers, was particularly high for certain diseases, e.g. cancer of the lung (10.8), bronchitis and emphysema (6.1), etc. For coronary artery disease the mortality ratio was 1.7. It seems that for effectively all causes of death (even accidents, suicide, and violence) the cigarette smokers had the higher death rate. Though cancer of the lung had the highest mortality ratio "coronary artery disease is the chief contributor to the excess number of deaths

of cigarette smokers over non-smokers" (page 113). This chapter makes no explicit claim that cigarette smoking is the cause of the mortality ratios being greater than 1.

Chapter 8 reaches its conclusions solely on the basis of seven prospective studies, presumably (though nothing is stated) for the reason that its authors were aware of the hazards in attempting to make rigorous inferences from retrospective studies (possible biases in retrospective studies are discussed in Chapter 9, pages 180–181). The main difficulty, it seems to me, is that it is very difficult to define accurately the population from which the hospital sample of, e.g., cancer patients has been drawn, and even more difficult to draw a sample of controls from that population.

In some of the prospective studies a clearly defined population existed, e.g., policy holders of U. S. Government Life Insurance Policies, and the intention was to take a 100 per cent sample of this population, but in general for these studies non-response averaged 32 per cent. Biases in the non-respondents could play havoc with inferences based on the respondents. Appendix I to Chapter 8 considers this, and on various assumptions about the behavior of the non-respondents concludes that such biases could, under rather extreme assumptions, make an actual mortality ratio of, for example, 4.0 be observed as 5.0, or an actual 7.0 be observed as 10.0. The possible effects of the non-respondents are tricky, however, as Doll found (page 97) that for British doctors the non-respondents had a higher death rate and relatively more smokers than the respondents. The cross-classification is, however, not given.

In other prospective studies, e.g. the 1963 Hammond study, there is no clearly defined population and hence the concept of percentage response is impossible to evaluate. Bias could exist in the recruitment into the sample, but it is stated (page 181) that its effects, if any, should decrease with time, and this is not observed. Mainland and Herrera [7] had previously wondered about the effects of biases in studies of this type, but the Report cites the gradual increase in the mortality ratio with time as weakening their criticism. The 1963 Hammond study specifically excluded "persons too ill to answer a questionnaire" (also illiterates, and persons who could not have been traced). Presumably these exclusions are responsible for the age adjusted death rate of the Hammond sample being substantially below that of U. S. males generally. In fact, for all seven prospective studies the samples are healthier than U. S. males (page 95). This may be largely due to the probable fact that the surveys are drawn from relatively high socio-economic groups. In specific surveys other reasons may also operate.

Even though cancer contributes only 26 per cent to the total excess number of deaths of cigarette smokers, it receives the most attention in the Report, as Chapter 9 being 136 pages is substantially the largest. Lung cancer, which contributes 16 per cent to the total number of excess deaths, receives the greatest emphasis and we will confine our attention accordingly.

Chapter 9 first reviews age-adjusted mortality rates for various countries, for various times (1900–1960), by site, by sex, income class, occupation, and ethnic group. It then discusses various carcinogens, some of which are present in tobacco smoke. Then follows pages 149–196 on lung cancer, reaching the

conclusion (page 199) "Cigarette smoking is causally related to lung cancer in men." We will now summarize the arguments that led to that conclusion.

Pages 150–161 review 29 retrospective studies and, briefly, the 7 prospective studies that had been discussed in Chapter 8. There follows a discussion of attempts to induce lung cancer with tobacco extracts and tobacco smoke in experimental animals. These have not been successful. However, industrial carcinogens were effective. Another section (page 167) reports that "genetic factors exert a determining influence on the spontaneous development and induction of lung tumors in mice."

Pages 175–176 discuss the correlation of national crude male death rates for lung cancer with per capita consumption of cigarettes. Using a 20 year lag, Doll obtained $r = 0.73$ with 11 countries. The "population" from which these 11 countries were "sampled" is not described.

Pages 179–182 discuss possible biases being responsible for the observed association, and concludes that the observed association is real.

The Report then moves to establish the causal significance of the association, its arguments following the five criteria that Chapter 3 listed as being part (or all?) of the "epidemiologic method."

(a) The first criterion is "consistency." The 29 retrospective and the 7 prospective studies all demonstrate the same association.

(b) "The most direct measure of the strength of the association between smoking and lung cancer is the ratio of lung cancer rates for smokers to non-smokers . . . " It is high, of the order of 9 to 1. Also "Important to the strength as well as to the coherence of the association is the dose-effect phenomenon. In every prospective study that provided this information, the dose-effect was apparent . . . "

(c) The specificity of the association "implies the precision with which one component of an associated pair can be utilized to predict the occurrence of the other." This seems a curious use of the word "specificity." A more satisfactory definition would seem to be Yerushalmy and Palmer's [15] "The basic assumption of such a test for specificity is that if the characteristic is not related to the disease in a causal way, then the relationship should not be restricted to the disease under study but should also be present with other disease entities. If the characteristic can be shown to be related only or mostly to the disease under study and not to many other disease entities, then our confidence that it is a cause-carrying vector for that disease is greatly increased."

This section of the Report discusses both these interpretations, its own and Yerushalmy and Palmer's, of "specificity." On the first, it comments that with diseases with multiple causes one cannot expect high specificity in this sense.

On the second, the Report states "The number of diseases in which the ratios remain significantly high, after consideration of the non-response bias, is not great enough to cast serious doubt on the causal hypothesis."

It further comments that even a single substance might cause several diseases and a mixture of substances, such as occur in tobacco smoke, might "produce more than a single disease." The Report concludes "Thus, it is reasonable to conclude that the association between cigarette smoking and lung cancer has a high degree of specificity."

(d) The next section, "Temporal Relationship of Associated Variables," makes the point that people smoke and then have cancer, rather than vice versa. The point of course, is that if *B* occurs later in time than *A* then it is difficult to maintain that *B* is a cause of *A*. This is obviously a necessary condition, but not sufficient.

(e) The final criterion is the "Coherence of the Association" which covers a number of points as follows.

(1) The historical increase in cigarette consumption is correlated with the historical increase in lung cancer.

(2) The sex difference in smoking is in the same direction as the difference in lung cancer rates.

(3) Page 186 reports that "although adjustment for smoking history does not equalize the urban-rural lung cancer mortality ratio, control on the urban-rural residence factor nevertheless leaves a large mortality risk difference between smokers and non-smokers." This seems to imply that correlation of residence with smoking cannot account for the observed association of lung cancer rate with smoking, and, vice versa, correlation of smoking with residence cannot account for the observed association of lung cancer with residence.

(4) A section on socio-economic differentials in lung cancer mortality remarks (page 187) "it will be nevertheless noted that the professional and farmer and farm manager groups had higher proportions of non-smokers among them than did the laborers and craftsmen. This finding is in the proper direction for compatability with the socio-economic differential in lung cancer mortality but the disparity does not appear to be sufficient to provide a satisfying correction."

(5) This section comments on the dose-response relationship and refers to the muddled position with respect to inhalation.

A section on the histopathologic evidence (page 189) refers mainly to work by Auerbach on epithelial changes in the trachea and bronchi:

"These changes were rarely seen among non-smokers, but increased in frequency and intensity with the number of cigarettes smoked daily by individuals without cancer and were most frequent and intense in patients dying of lung cancer." The opinion is expressed (page 172) "it seems probably that some of the lesions found in the tracheobronchial tree in cigarette smokers are capable of developing into lung cancer."

It is difficult for one not a specialist in this area to assess the significance of these observations. The Report concludes (page 189) "Thus, the histopathologic evidence derived from laboratory and clinical material support the cigarette smoking—lung cancer hypothesis" so the Committee does not appear to regard this time as crucial. Presumably having established to its satisfaction that smoking causes lung cancer since the criteria of the "epidemiologic method" are allegedly satisfied, the Report in pages 190–193 discusses the "Constitutional Hypothesis," i.e. "the alternative hypothesis that both smoking of cigarettes and cancer of the lung have a common cause . . ." and concludes "that genetic factors play a minor role" (page 192). The main

reasons for this conclusion appear to be (a) the presumption that the genetic hypothesis would have to be very complicated to include the dose-response relationship, etc., and (b) the rapid historical rise of lung cancer.

Finally, sundry factors such as occupational hazards, urbanization, air pollution, etc., are mentioned, and on page 196 the conclusion of causality is unequivocally stated.

This review has so far been largely restricted to a summarization of the relevant arguments leading to the above important conclusion. I will now discuss the question of whether the available evidence justifies the Report in reaching that conclusion. It is possible, of course, for the conclusion to be in fact correct but for the evidence in its favor to be inadequate. We are all familiar with the accused who has to be found not guilty, because of insufficient evidence, though we are personally certain of his guilt. Let us now comment on various parts of the Report's position.

In my opinion a key factor in determining one's opinions is one's interpretation of Table 26, pages 109–110, which shows the mortality ratios, individually by separate study and also jointly, for 25 causes of death, ordered from cancer of lung with a total mortality ratio of 10.8, bronchitis and emphysema with a total mortality ratio of 6.1, down to cancer of the rectum (1.0) and cancer of the intestines (0.9). The latter cause of death is the only one out of the 25 to show a total mortality ratio of less than 1. The total over all studies column of Table 26 is reproduced here as Table 1. The median of the 25 total mortality ratios is 1.5, and 8 are greater than 2.0. One could adopt an extreme null hypothesis attitude and admit that only those causes of death for which the total mortality ratio was statistically significantly greater than 1 are genuinely associated with cigarette smoking. The technical details of this statistical problem would be difficult on account of the various "sampling" procedures involved, allowances for possible biases and dependencies, the varying sample sizes, and the multiplicity of the tests being made. If we take the twenty-first cause of death, "accidents, suicides, and violence," with a total mortality ratio of 1.2, it is noteworthy that in each of the seven separate studies the mortality ratio is greater than 1. In my opinion virtually every tabulated cause of death shows a mortality ratio greater than 1. I do not regard even the twenty-fifth cause of death, cancer of the intestines, as significantly an exception to this, as if we look at estimates of 25 parameters the smallest of the 25 is likely to be smaller than its population value, and at 0.9 it would not take much of a sampling error to bring the ratio over 1. However, whether one feels that probably all 25 causes of death have mortality ratios greater than 1, or merely the first 21 out of the 25, or the first 15, makes little difference in my opinion. In the discussion I give below, I shall use the phrase "virtually all causes of death" but it could be replaced by some such phrase as "21 out of 25 causes of death" without changing appreciably the import of the discussion.

One's interpretation of Table 26 determines whether one accepts the Report's conclusion (page 185) "Thus, it is reasonable to conclude that the association between cigarette smoking and lung cancer has a high degree of specificity" as, on the one hand, reasonable, or on the other hand, as an extraordinary statement completely in violation with the facts. The Report

TABLE I. NUMBERS OF EXPECTED AND OBSERVED DEATH FOR SMOKERS OF CIGARETTES ONLY, AND MORTALITY RATIOS, TOTALLED OVER ALL SEVEN PROSPECTIVE STUDIES. TAKEN FROM TABLE 26, PAGE 110, OF THE REPORT

Cause of death	Deaths		Mortality ratio
	Expected	Observed	
Cancer of lung	170.3	1,833	10.8
Bronchitis, emphysema	89.5	546	6.1
Cancer of larynx	14.0	75	5.4
Cancer of oral cavity	37.0	152	4.1
Cancer of esophagus	33.7	113	3.4
Stomach and duodenal ulcers	105.1	294	2.8
Other circulatory diseases	254.0	649	2.6
Cirrhosis of liver	169.2	379	2.2
Cancer of bladder	111.6	216	1.9
Coronary artery disease	6,430.7	11,177	1.7
Other heart diseases	526.0	868	1.7
Hypertensive heart disease	409.2	631	1.5
General arteriosclerosis	210.7	310	1.5
Cancer of kidney	79.0	120	1.5
All other cancer	1,061.4	1,524	1.4
Cancer of stomach	285.2	413	1.4
Influenza, pneumonia	303.2	415	1.4
All other causes	1,508.7	1,946	1.3
Cerebral vascular lesions	1,461.8	1,844	1.3
Cancer of prostate	253.0	318	1.3
Accidents, suicides, violence	1,063.2	1,310	1.2
Nephritis	156.4	173	1.1
Rheumatic heart disease	290.6	309	1.1
Cancer of rectum	207.8	213	1.0
Cancer of intestines	422.6	395	.9
All causes	15,653.9	26,223	1.68

accepts "specificity" as one of the criteria of "the epidemiologic method" (pages 20, 183), and in my opinion the way it claims the facts are in conformity with the criterion is to flatly ignore the facts.

In my opinion, therefore, Table 26 raises difficulties which the Report totally fails to face up to. The difficulties are as follows.

If one believes that the observed association between smoking and lung cancer is substantially real, and not an artefact of biased sampling, then one would seem required also to accept the observed association between smoking and almost all causes of death as substantially real.

If one believes that the observed association between cigarette smoking and virtually all causes of death is substantially real, then one must take one of the following positions:

(a) All the associations are due to causation.

(b) Some of the associations are due to causation and others to correlations, concealed or otherwise.

(c) All of the associations are due to correlations, concealed or otherwise.

Positions (a) and (c) may present the fewest difficulties. If one adopts position (a), then one is under some obligation to provide hypotheses as to possible mechanisms, or to hold out hope that future research will provide these hypotheses. The difficulty with the smoking hypothesis is that it has not really got to first base on even lung cancer, let alone the other 24 causes of death, even though the matter has been under intensive investigation for ten years or more.

The most potent carcinogen identified in tobacco smoke is benz (a) pyrene (page 27) and "it is present in much larger quantity than any of the other carcinogens listed." Cigar smoke has almost 4 times as much benz (a) pyreme, and pipe smoke about 10 times as much as cigarette smoke (page 58), yet pipes and cigars are pretty well innocent of the charge of association with lung cancer. (However, the Committee does conclude that the smoking of pipes is causally related to cancer of the lip (page 204).)

Apart from the above specific measurements on smoke, the difference between cigarettes and cigars (and pipes) is a puzzling feature of the indictment of tobacco. It could be, of course, that the variety of tobacco used differs significantly, that the tobaccos are cured differently, that cigars and pipes do not use cigarette paper, that possibly pipes and cigars burn at lower temperatures, or that many cigarette smokers inhale whereas few pipe or cigar smokers do. But if inhalation is the crucial item, then it should show up strongly when cigarette smokers are analyzed into inhalers and non-inhalers, and as reported above, the present evidence on this point is not clear.

As stated earlier, animal experiments to demonstrate carcinogenic effects of tobacco smoke have proved unsuccessful.

While hypotheses for a causative effect for lung cancer are still unsatisfactory after a decade of research, the situation is even emptier for other diseases.

The degree of one's belief in the plausibility of the smoking hypothesis thus is considerably affected by the extent to which one thinks it probably that satisfactory mechanisms will be provided to account for all, or most, of the diseases.

If one adopts position (b), then one is in the unenviable position of admitting that concealed correlations account for diseases D_i, $i=k, k+1, \cdots, n$, but are not responsible for diseases D_i, $i=1, 2, \cdots, k-1$. It seems to me that once one admits that concealed correlations account for a substantial number of the observed associations, then one has to work very hard to disprove the hypothesis that they may account for all the observed correlations.

The Report uses the historical correlation of rise in lung cancer mortality with increase in per capita consumption of cigarettes as one factor to support the coherence of the association (page 185). All statisticians know that the presence of a positive, zero, or negative correlation between two variables observed over time has been the basis of more ludicrous nonsense than any other statistical procedure. For example, the incidence of cancer of the stomach has been declining for many years (Dorn and Cutler [18]), but only a madman would infer that the increased smoking has caused the decreased stomach cancer. Incidentally, it is surprising that the Report makes no mention of this fact in the section on stomach cancer (pages 225–229) though it is mentioned

on page 135. The Report omits any comment on the weakness of this item in its discussion of the "Coherence of the Association" on page 185.

Similar considerations apply to spatial correlations, which are presented on page 176.

The fact that the hypothesis that cigarette smoking is a cause of cancer appears by and large to be in conformity with the data, and hence that this hypothesis is acceptable, does not rule out the possibility that there are other hypotheses also in conformity with the data. As Sir Ronald Fisher [19] wrote almost thirty years ago:

"For the logical fallacy of believing that a hypothesis has been proved to be true, merely because it is not contradicted by the available facts, has no more right to insinuate itself in statistical than in other kinds of scientific reasoning." And as Yule [20], quoted by Irwin [21], wrote about forty years ago: " 'You can prove anything by statistics' is a common gibe. Its contrary is more nearly true—you can never prove anything by statistics. The statistician is dealing with the most complex cases of multiple causation. He may show that the facts are in accordance with this hypothesis or that. But it is quite another thing to show that all other possible hypotheses are excluded, and that the facts do not admit of any other interpretation than the particular one he may have in mind."

It is not clear to me what are the optimal, or even satisfactory, procedures for inference. In the case where there are two competing hypotheses, one might estimate the "plausibility ratio," analogous to the likelihood ratio, and if it is very small, $\ll 1$, or very large, $\gg 1$, one can reach a conclusion. If the plausibility ratio is in the neighborhood of 1, then no decision can be reached. One would further wish that if the more plausible hypothesis is chosen as a null hypothesis, then a test of this null hypothesis can be accepted at a large P value, for otherwise one would suspect that neither hypothesis was correct, and if the less plausible hypothesis is tested as a null hypothesis one would want to it be rejected at a small P value.

The main alternative to the smoking causes-cancer hypothesis is the genetic hypothesis, and there are several odd pieces of information that give plausibility to it.

(1) There is some evidence that non-smokers, cigarette smokers, and pipe and cigar smokers are morphologically different (Chapter 15, reference to Seltzer [17]).

(2) There is some evidence that the various classes of smokers and non-smokers are psychologically different (Chapter 14, reference to Eysenck [16]).

(3) There is some evidence that identical twins are more alike than non-identical twins in their smoking habits (page 190, references to Fisher [22], Friberg et al [13], and Raaschou-Nielsen [24]).

(4) "Foreign-born migrants to the United States as a group have age-adjusted death rates for cancer of the esophagus and stomach about twice those recorded for native-born white males and females," etc., (page 134).

(5) "The several ethnic groups in the United States display their own characteristic patterns of successes and deficits in risk by site" of cancer (page 135).

It is true, of course, that these phenomena ((4) and (5)) could be the result

of social, dietary, or other customs imported by the immigrants and maintained for one or more generations, or could be a confounding of the effects of the environment of the region of the U. S. in which each ethnic group tended to settle. Thus these phenomena can be interpreted in terms of environmental effects rather than genetic effects.

(6) Yerushalmy [14] presented data strongly suggesting that whether or not the husband smokes is associated with the incidence of premature births.

(7) Hammond [13] gives results showing marked association of death rate with longevity of parents and grandparents.

On my reading, the main reasons the Report gives for rejecting the genetic hypothesis are (1) the necessary complexity of the genetic hypothesis, and (2) the unlikelihood that the genetic pool can have changed sufficiently in the past fifty years to account for the historical rise in lung cancer rate.

For the first difficulty, different genotypes need to be hypothesized for the various classes such as cigarette smokers, cigarette smokers who give up cigarette smoking, cigar smokers, etc. It seems to me that an enormous number of traits are transmitted genetically, from color of skin and tendency to diabetes and tendency to baldness, and it is completely plausible that the tendency to be a cigarette smoker in varying degrees, a pipe or cigar smoker, and so on can be carried by the genetic code.

The second objection referred to above seems to me to overlook a possible fallacy in the interpretation of historical change in death rates. Firstly, possibly part of the long term apparent rise may be an illusion caused by fashions in diagnosis and by improvements in diagnostic techniques. It is quite probable that several decades ago many cases of lung cancer would have been diagnosed as tuberculosis. The Report claims (page 140) that over the shorter term, from 1947 on, particularly for data for Connecticut and New York, the increase in genuine as there have been "no significant advances in diagnostic methods" and in these regions "a high percentage of the cases reported have microscopic confirmation." This comment does not quite bear on the point at issue. The suggestion of the skeptics is not that cases of tuberculosis are now being falsely diagnosed as lung cancer but on the contrary that in the past cases of lung cancer were falsely diagnosed as tuberculosis.

The second argument is more complex. The generation born in 1880 reached the age of 20 in 1900, and a substantial fraction, 24 per cent, had died by this time, presumably largely due to the traditional infectious diseases of childhood. The generation born in 1900 reached the age of 20 in 1920, and a lesser fraction of this generation had died by that age, namely 15 per cent. Therefore, the 1900 cohort at the age of 50 represents quite a different stratum from the 1880 cohort at the same age of 50, since the former includes the "weaklings" who were eliminated from the latter. The relatively large number of cases of lung cancer observed in the 1900 cohort can merely be largely those who would not have survived to run the risk of lung cancer if they had been born 20 years earlier. On this model the alleged historical increase in rate of lung cancer can be readily accounted for. The model is speculative, however.

Rather strangely, the Report does not include in the section devoted to refuting the constitutional hypothesis (pages 190–193) the observations by Corn-

TABLE 2

	X		X^c		Sums over X and X^c	
	S	S^c	S	S^c	S	S^c
C	0.00450	0.00050	0.00005	0.00045	0.00455	0.00095
C^c	0.44550	0.04950	0.04995	0.44955	0.49545	0.49905
Sums	0.45000	0.05000	0.05000	0.45000		

field et al [12] that if the mortality ratio for smokers to non-smokers for a particular disease is k, then factor X must be present at least k times more frequently among smokers than among non-smokers. The Report does refer to this result, without citation of source, in the section on specificity (page 184), where its relevance escapes me. It is clear, however, that Cornfield et al regard this result as one of the key arguments against the constitutional hypothesis, as they cite it prominently in the Summary to their paper [12]:

"The magnitude of the excess lung cancer risk among cigarette smokers is so great that the results can not be interpreted as arising from an indirect association of cigarette smoking with some other agent or characteristic, since this hypothetical agent would have to be at least as strongly associated with lung cancer as cigarette use; no such agent has been found or suggested."

A simple arithmetical illustration of this (mathematical) phenomenon is developed below. Suppose that the population is divided 50:50 into X and X^c. Suppose that amongst X people the probability of being a smoker is 0.9 and the probability of getting cancer is 0.01, and that these probabilities are independent. Suppose further that among X^c people the probability of being a smoker is 0.1 and the probability of getting cancer is 0.001, and that these probabilities are independent. Then the resulting probabilities are in Table 2. The cancer rate among smokers is $0.00455/(0.00455+0.49545)=0.0091$ and the cancer rate among non-smokers is $0.00095/(0.00095+0.49905)=0.0019$: the mortality ratio for cancer for smokers/non-smokers is thus $0.0091/0.0019=4.79$. The X rate among smokers is $0.45/(0.45+0.05)=0.90$ and among non-smokers is $0.05/(0.05+0.45)=0.10$; the ratio for X for smokers/non-smokers is thus $0.90/0.10=9.0$. This illustrates Cornfield et al's point that the X ratio for smokers/non-smokers must be greater than the mortality ratio for smokers/non-smokers. I do not feel that the numerical values of the parameters inserted in this simple model are ridiculous, yet with independence of S and C for X and for X^c we came out with a large mortality ratio for the population as a whole. I do not see why the fact that the X ratio is larger than the mortality ratio is any embarrassment to the constitutional hypothesis.

The Report quotes various results from Hammond's latest report [13] which illustrates various phenomena which can be attributed to various constitutional hypotheses, though one could also easily construct other causal hypotheses in each case. The "effect" of the use of tranquilizers, for example is shown in Table 3. (All numbers quoted in Tables 3–6 are age standardized death rates from all causes for males.) Recalling that this data was collected by volunteers,

TABLE 3

Tranquilizers	Never smoked regularly	Cigarettes 20+ a day
Do not use	755	1501
Use	1308	2286

TABLE 4

Fried Food	Never smoked regularly	Cigarettes 20+ a day
No Fried Food Eaten	1208	2573
Fried Food 3-4 times a wk.	642	1714

often in a "social" environment, it is probable that some tranquilizer users did not admit to this habit, and therefore appear in the "do not use" row, hence raising the death rate for this row. The real "effect" of tranquilizers is therefore probably even more dramatic than the table shows.

Tables 4, 5, 6 are from Hammond; the first two are not quoted in the Report. Hammond's data also show the well known, but as far as I know, never satisfactorily explained, differences in death rates between married, single, widowed, and divorced persons. It also shows a marked hereditary effect, a marked height effect, and a moderate education effect. Hammond does not give cross-classifications, other than on smoking.

It is noteworthy that in all the above two-way cross-classifications, and others in Hammond's report not quoted here, the smoking category always has

TABLE 5

Sleep	Never smoked regularly	Cigarettes 20+ a day
<5 hrs.	2029	3936
5 hrs.	1121	2655
6 hrs.	805	1601
7 hrs.	626	1426
8 hrs.	813	1562
9 hrs.	967	1729
10+hrs.	1898	2694

TABLE 6

Exercise	Never smoked regularly	Cigarettes 20+ a day
None	834	1416
Slight	579	1347
Moderate	486	1065
Heavy	474	998

the higher death rate. The Report mentions an unpublished analysis supplied to the Committee by Ipsen and Pfaelzer, with the conclusion, "Further, the correlation of any of these variables with cigarette smoking was too weak to reduce markedly the correlation of cigarette smoking with mortality after adjustment for the other variable." It is a pity that the Report does not give any detail on this analysis, as it may be that this analysis would have carried the greatest conviction. However, this is speculation, and we can only judge the Report by what it contains and by its references to the published literature.

One can interpret the results of Tables 3–6 in various ways. The association with longevity of parents and grandparents probably has a genetic interpretation, though it is possible to interpret it environmentally: it could be that the grandparents settled in an environment or adopted habits and customs that were favorable to longevity, and which they passed on to the later generations. The association with tranquilizers could be interpreted as causal, as is the association with cigarettes, but it appears equally plausible to hypothesize that the need for tranquilizers, and/or cigarettes, is merely a reflection of an underlying facet of the human organisms involved. Likewise, the associations with fried food eating, sleep, and exercise can be interpreted either as causal or as concealed correlations.

It will be recalled that the Committee's assignment was "if possible, to reach some definitive conclusions on the relationship between smoking and health in general" (page v).

"This committee would produce and submit to the Surgeon General a technical report containing evaluations and conclusions" (page 8).

The Committee's "Judgment in Brief" was "Cigarette smoking is a health hazard of sufficient importance in the United States to warrant appropriate remedial action" (page 33).

The detailed conclusions cover pages 37–40; the most important are "Cigarette smoking is causally related to lung cancer in men."

"Male cigarette smokers have a higher death rate from coronary artery disease than non-smoking males, but it is not clear that the association has causal significance."

My opinion is that the Committee has not established the case for causality in lung cancer. My reasons for this opinion are, to recapitulate;

(a) This conclusion can only be justified by proving the genetic hypothesis to be false, and this the Committee has failed to do.

(b) Even if there was no competing hypothesis, the case for causality is at the present time significantly weakened by the gross nonspecificity of the association and by the absence of hypothetical physical-chemical mechanisms.

My opinion that the Committee has failed to offer a satisfactory proof of the hypothesis of causality does not imply that I believe that the causality hypothesis has been proved false. It implies nothing more than I believe that it is not possible to reach definitive conclusions at this time.

Although "Recommendations for actions were not to be part of the Phase I committee's responsibility" (page 8), yet as noted above the Committee did conclude that "appropriate remedial action" was warranted. I take this as an excuse to append some personal reflections.

Passing over the question of whether any governmental action is warranted, each individual cigarette smoker has to decide whether to make the effort to give up cigarette smoking. To make this decision he would need to know (a) the change in his expectation of life assuming that lung cancer alone is causally connected with cigarette smoking, (b) the change in his expectation of life assuming that virtually all causes of death are causally connected with cagarette smoking.

The Report does not appear to provide these figures (searching in the Report for any particular topic is sometimes difficult as the Report has no index), which is rather surprising in view of the fact that an assessment of the *magnitude* of the health hazard was one of the Committee's prime responsibilities (page 8).

The individual cigarette smoker would then have to weight the above changes in expectation by his assessment of the question of whether the cigarette smoking was a cause. He would have to relate this expected change in expectation of life to the analogous probable expected changes in expectation of life caused by various other of his activities which probably cause decreases in expectation of life, such as over-eating, undersleeping, under-exercising, travelling by automobile, etc. He would also have to weigh the possibility of such adverse consequences of giving up smoking as in due course becoming appreciably overweight. He would have to assign utilities to the expected change in expectation of life and to the pleasure and satisfaction he gets from cigarette smoking.

It is unfortunate that the Report does not give a table of changes in expectation of life, for if the change at, say, age 30 was 10 years, then even if one assigned a probability of 1/5 to the cigarette smoking causality hypothesis being true, the net change of 2 years would give one pause to consider, whereas if the change at age 30 is only 2 years, then the net change of 0.4 years on an average expectation of about 40 years might not be taken too seriously by some people.

REFERENCES

[1] Müller, F. J., "Tabakmissbrauch und Lungencarcinom," *Zeitschrift fur Krebsforschungen*, 49 (1939), 57–84.

[2] Doll, R., and Hill, A. B., "A study of the aetiology of carcinoma of the lung," *British Medical Journal*, 2 (1952), 1271–86.

[3] Doll, R., and Hill, A. B., "The mortality of doctors in relation to their smoking habits; a preliminary report," *British Medical Journal*, 1 (1954), 1451–5.

[4] Hammond, E. C., and Horn, D., "The relationship between human smoking habits and death rates: A follow-up study of 187,766 men," *Journal of the American Medical Association*, 155 (1954), 1316–28.

[5] Neyman, J., "Statistics—servant of all sciences," *Science*, 122 (1955), 401–6.

[6] Berkson, J., "The statistical study of association between smoking and lung cancer," *Proceedings of the Staff Meetings of the Mayo Clinic*, 30 (1955), 319–48.

[7] Mainland, D., and Herrera, L., "The risk of biased selection in forward-going surveys with nonprofessional interviewers," *Journal of Chronic Diseases*, 4 (1956), 240–4.

[8] Cutler, Sidney J., "A review of the statistical evidence on the association between smoking and lung cancer," *Journal of the American Statistical Association*, 50 (1955), 267–82.

[9] Fisher, R. A., Letter to the Editor, *British Medical Journal*, 2 (1957), 297–8.

[10] Berkson, J., "Smoking and lung cancer: some observations on two recent reports," *Journal of the American Statistical Association*, 53 (1958), 28–38.
[11] Berkson, J., "Smoking and lung cancer," *American Statistician*, 17 (1963), 15–22.
[12] Cornfield, Jerome, et al., "Smoking and lung cancer: Recent evidence and a discussion of some questions," *Journal of the National Cancer Institute*, 22 (1959), 173–203.
[13] Hammond, E. Cuyler, "Smoking in relation to mortality and morbidity: Findings in first 34 months of follow-up in a prospective study started in 1959." Statistical Research Section, Medical Affairs, Department of the American Cancer Society.
[14] Yerushalmy, J., "Statistical considerations and evaluation of epidemiological evidence," in James, George, and Rosenthal, Theodore, *Tobacco and Health*. Springfield: Charles C Thomas, 1962, 208–30.
[15] Yerushalmy, J., and Palmer, Carroll E., "On the methodology of investigations of etiologic factors in chronic diseases," *Journal of Chronic Diseases*, 10 (1959), 27–40.
[16] Eysenck, H. J., Tarrant, M., Woolf, M., and England, L., "Smoking and personality," *British Medical Journal*, 1 (1960), 1456–60.
[17] Seltzer, Carl C., "Morphologic constitution and smoking," *Journal of the American Medical Association*, 183 (1963), 639–45.
[18] Dorn, Harold F., and Cutler, Sidney J., "Morbidity from cancer in the United States," Public Health Monograph No. 56, 1959.
[19] Fisher, R. A., Letter to the Editor, *Nature*, 136 (1935), 474.
[20] Yule, G. V., "The function of scientific method in scientific investigation," Industrial Fatigue Research Board Report, 28 (1924).
[21] Irwin, J. O., "The place of mathematics in medical and biological statistics," *Journal of the Royal Statistical Society, Series A*, 126 (1963), 1–45.
[22] Fisher, R. A., "Cancer and smoking," Letter to the Editor, *Nature*, 182 (1958), 596.
[23] Friberg, L., Kaij, L., Dencker, S. J., and Tonsson, E., "Smoking habits in monozygotic and dizygotic twins," *British Medical Journal*, 1 (1959), 1090–2.